



Jurnal Edik Informatika
Penelitian Bidang Komputer Sains dan Pendidikan Informatika
V5.i1(41-53) 2018

ISSN : 2407-0491
E-ISSN : 2541-3716

METODE DATA MINING ASSOCIATION RULE DENGAN ALGORITMA FP-GROWTH UNTUK MENGETAHUI KELULUSAN TEPAT WAKTU MAHASISWA (STUDI KASUS STKIP PGRI SUMATERA BARAT)

Satrio Junaidi¹, Thomson Mary²

^{1,2}Dosen Pendidikan Informatika STKIP PGRI Sumbar

^{1,2}Jl. Gunung Pangilun No. 02, Padang

satriojunaidy@gmail.com, Thomsonmary1980@gmail.com

<http://dx.doi.org/10.22202/ei.2018.v5il.3427>

Abstrak

Sekolah Tinggi Keguruan dan Ilmu Pendidikan (STKIP) PGRI Sumatera Barat memiliki 14 program studi S1. Salah satunya adalah Program Studi Pendidikan Matematika STKIP PGRI Sumatera Barat. Setiap tahun jumlah mahasiswa yang lulus mencapai angka ratusan bahkan mencapai ribuan. Maka dari itu Prodi Pendidikan Matematika STKIP PGRI Sumatera Barat butuh media mengevaluasi untuk akademik dari mahasiswa lulusan yang dibutuhkan sebagai bahan pertimbangan untuk menentukan mahasiswa yang lulus tepat waktu. Bertujuan untuk membantu pihak Prodi Pendidikan Matematika Sumatera Barat untuk mengetahui jumlah kelulusan mahasiswa tepat waktu setiap tahun. Data yang diolah adalah data akademik mahasiswa yang dilihat dari IPK dan Total SKS, kemudian untuk mengetahui jumlah kelulusan mahasiswa tepat waktu menggunakan metode Data Mining Association Rule dengan algoritma FP-Growth dan juga menggunakan aplikasi WEKA (Weikato Environment Knowledge and Analisis). Hasil didapatkan bahwa rule terbaik dijadikan patokan untuk mengetahui mahasiswa lulus tepat waktu dengan support 14,8 % dan confidence 93,7 % . hasil support dan confidence terendah diprediksi mahasiswa tidak lulus tepat waktu.

Kata Kunci : *Tingkat kelulusan mahasiswa; Association Rule; Algoritma FP-Growth*

ABSTRACT

The College of Teacher Training and Education (STKIP) PGRI West Sumatra has 14 S1 courses. One of them is Mathematics Education Studies Program STKIP PGRI West Sumatra. Every year the number of graduating students reaches hundreds even thousands. Therefore, the Mathematics Education Study Program of STKIP PGRI West Sumatra needs the media to evaluate for the academic of the graduate students needed for consideration to determine the students who pass on time. Aims to help the Mathematics Education Study Program of West Sumatra to find out the number of graduation students on time each year. Processed data is

academic data of students viewed from the GPA and Total SKS, then to know the number of students pass on time using Data Mining Association Rule method with FP-Growth algorithm and also using WEKA (Weikato Environment Knowledge and Analipsis) application. The results obtained that the best rule used as a benchmark to know the students graduated on time with support 14.8% and confidence of 93.7%. The lowest support and confidence results predicted that students do not pass on time.

Keywords: *The Graduation rate of students; Association Rule; FP-Growth Algorithm*

1.PENDAHULUAN

Teknologi informasi dalam jangka waktu yang sesingkat-singkatnya berkembang dengan cepat. Penyedia informasi menjadi sarana penting untuk menganalisa setiap data yang ada untuk mendapatkan bahan pertimbangan dari informasi yang tersedia. Perguruan Tinggi adalah tempat pendidikan bagi seseorang yang telah selesai di sekolah menengah atas. Mahasiswa merupakan aset yang sangat penting disebuah perguruan tinggi. maka, faktor yang menentukan kualitas perguruan tinggi adalah persentasi kemampuan mahasiswa untuk menyelesaikan studi tepat waktu (Prasetyo T.F, et al., 2016). Lulus tepat waktu merupakan hal yang sangat penting disebuah perguruan tinggi karena tingkat kelulusan sebagai dasar efektifnya suatu kelembagaan. Jika terjadi penurunan tingkat kelulusan maka akan menjadi suatu permasalahan yang akan mempengaruhi akreditasi di sebuah perguruan tinggi.

Di dalam peraturan akademik STKIP PGRI Sumatera Barat tahun 2013 pada BAB IV pasal 10 ayat 1 di sebutkan bahwa "Setiap Program Studi pada jenjang pendidikan mempunyai beban dan lama studi sebagai berikut: S1 =144-160 sks dengan lama studi 8-14 semester" (Peraturan Akademik, 2013). Program Sarjana (S1) reguler yang memiliki beban studi sekurang-kurangnya 144 (seratus empat puluh empat) sks (satuan kredit semester) dan sebanyak-banyaknya 160 (seratus enam puluh) sks

yang dijadwalkan untuk 8 (delapan) semester dan paling lama 14 (empat belas) semester. Pada tahun 2017 wisuda 55 prodi pendidikan matematika diketahui sedikitnya lulusan tepat waktu. Maka dari itu yang menjadi permasalahannya adalah tidak seimbang jumlah mahasiswa yang dibandingkan dengan jumlah mahasiswa yang telah menyelesaikan masa studinya. Hal itu sangat berpengaruh terhadap proses akreditasi di Perguruan Tinggi.

Masalah tersebut sangat menarik untuk diteliti agar dapat mengetahui faktor apa yang menyebabkan hal itu sampai terjadi. Oleh karena itu, dengan adanya data induk mahasiswa dan data kelulusan mahasiswa, dapat diketahui informasi tingkat kelulusan mahasiswa melalui *Data Mining*. Penggunaan *Data Mining* diharapkan dapat memberikan pengetahuan-pengetahuan yang sebelumnya tersembunyi di dalam gudang data sehingga menjadi informasi yang diharapkan.

Data Mining didefinisikan sebagai proses menemukan pola-pola dalam data. Proses ini otomatis atau seringnya semi-otomatis. Pola yang ditemukan harus penuh arti dan pola tersebut memberikan keuntungan, biasanya keuntungan secara ekonomi. Data yang dibutuhkan dalam jumlah besar (Santoso H, et al., 2016). Dalam proses *Data Mining* penulis memerlukan beberapa perhitungan yang menggunakan algoritma. Algoritma yang digunakan untuk mempermudah proses perhitungan agar hasil

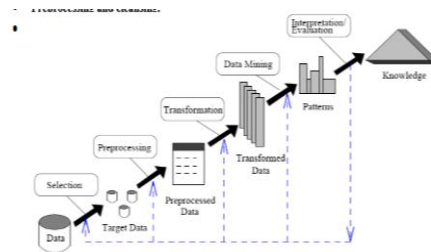
yang didapat tepat.

Berangkat dari permasalahan di atas maka untuk mengetahui kelulusan tepat waktu mahasiswa penulis menggunakan metode asosiasi *Data Mining* dengan pemilihan algoritma *FP-Growth*. *FP-Growth* adalah salah satu *alternative* algoritma yang dapat digunakan untuk menentukan himpunan data yang paling sering muncul (*frequent itemset*) dalam sebuah kumpulan data. Asosiasi adalah fungsi *Data Mining* yang ditemukan, korelasi antara item dalam kumpulan data yang besar. Aturan dihasilkan berdasarkan hubungan antara *item* yang terjadi (K. Sumathi, et al., 2016). Tugas asosiasi dalam *Data Mining* adalah menemukan atribut yang muncul dalam satu waktu (T ampubolon H, et al., 2013). *Association rule* adalah suatu metode *Data Mining* yang bertujuan untuk mencari sekumpulan items yang sering muncul bersamaan.

2. TINJAUAN PUSTAKA

Data Mining atau sering disebut sebagai Knowledge Discovery in Database (KDD) adalah kegiatan yang meliputi pengumpulan, pemakaian data historis untuk menemukan keteraturan, pola atau hubungan dalam data berukuran besar. Keluaran Data Mining ini bisa dipakai untuk membantu pengambilan keputusan di masa depan (Fadlina, 2014). Data Mining dan KDD sering digunakan secara bergantian karena Data Mining adalah bagian kunci dari proses KDD. Istilah Penemuan Pengetahuan dalam database atau KDD singkatnya, mengacu pada proses pencarian pengetahuan yang luas dalam data, dan menekankan penerapan metode penambangan data tingkat tinggi. Tujuan pemersatu proses KDD adalah mengekstrak pengetahuan dari data dalam konteks data yang besar (Priyadharsini & Thanamani A.S, et al., 2014). Hal ini dilakukan dengan menggunakan metode Data Mining (algoritma) untuk mengekstrak (mengidentifikasi) apa yang dianggap

pengetahuan, sesuai dengan spesifikasi ukuran dan ambang batas, menggunakan database bersamaan dengan preprocessing yang diperlukan, dan transformasi basis data tersebut



Gambar 2.1. Proses KDD

Knowledge Discovery in Database adalah suatu proses terdiri atas sembilan langkah langkah (Priyadharsini & Thanamani A.S, et al., 2014):

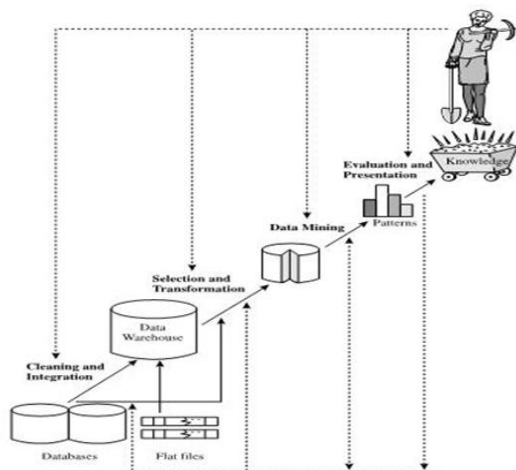
1. *Developing an understanding of the application domain* (Mengembangkan pemahaman tentang domain aplikasi).
2. *Selecting and creating a data set on which discovery will be performed* (Memilih dan membuat kumpulan data pada penemuan mana yang akan dilakukan).
3. *Preprocessing and cleansing* (Proses dan pembersihan).
4. *Data Mining (Data Mining)*.
5. *Choosing the appropriate Data Mining task* (Memilih tugas *Data Mining* yang tepat).
6. *Choosing the Data Mining algorithm* (Memilih algoritma *Data Mining*).
7. *Employing the Data Mining algorithm* (Menggunakan algoritma *Data Mining*).
8. *Evaluation* (Evaluasi).
9. *Using the discovered knowledge* (Menggunakan pengetahuan yang ditemukan)

2.1 Data Mining

Data Mining didefinisikan sebagai serangkaian proses untuk menggali nilai tambah dari suatu kumpulan data berupa pengetahuan yang selama ini tidak diketahui secara manual. *Data Mining* menganalisa hubungan dan pola dari data transaksi yang tersimpan berdasarkan kebutuhan *User* (Chachkamy M.H & Sadeghiyan B., 2013). *Data Mining* didefinisikan sebagai proses menemukan pola-pola dalam data. Proses ini otomatis atau seringnya semi-otomatis. Pola yang ditemukan harus penuh arti dan pola tersebut memberikan keuntungan (Santoso H, et al., 2016).

2.2 Tahapan-tahapan pada Data Mining

Karena *Data Mining* adalah suatu rangkaian proses, *Data Mining* dapat dibagi menjadi beberapa tahap. Tahap-tahap tersebut bersifat interaktif di mana pemakai terlibat langsung atau dengan perantara basis pengetahuan. Tahap-tahap ini diilustrasikan di gambar 2.4.



Gambar 2.4 Tahapan Data Mining

Tahapan *Data Mining* adalah sebagai berikut (Santoso H, et al., 2016):

1. Pembersihan data (*data cleaning*)

Pada tahap ini terjadi proses penghapusan atau menghilangkan *noise* dan data yang tidak konsisten atau data tidak relevan.

1. Integrasi data (*data integration*)

Pada tahap Integrasi data terjadi penggabungan data dari berbagai *database* ke dalam satu *database* baru.

2. Transformasi data (*data transformation*).

Pada transformasi data, data diubah atau digabung ke dalam format yang sesuai untuk diproses dalam *Data Mining*.

3. Proses Mining

Proses *Mining* merupakan proses utama yang dilakukan saat metode diterapkan untuk menemukan pengetahuan berharga dan tersembunyi dari data.

4. Evaluasi pola (*pattern evaluation*)

Untuk mengidentifikasi pola-pola menarik ke dalam *knowledge based* yang ditemukan.

5. Presentasi pengetahuan (*knowledge presentation*).

Merupakan visualisasi dan penyajian pengetahuan mengenai metode yang digunakan untuk memperoleh pengetahuan yang diperoleh pengguna.

2.3 Asosiasi

Asosiasi memiliki tugas dalam *Data Mining* yaitu untuk menemukan atribut yang muncul dalam satu waktu. aturan asosiasi (*Association Rules*) atau analisis afinitas (*affinity analysis*) berkaitan dengan studi tentang “apa bersama apa”. Salah satu implementasi dari asosiasi adalah *market basket analysis* atau analisis keranjang belanja, Contohnya pada studi transaksi di supermarket, misalnya seseorang yang membeli susu bayi juga membeli sabun mandi. Pada kasus ini berarti susu bayi bersama dengan sabun mandi. aturan asosiasi juga sering dinamakan *market basket analysis*, karena awalnya berasal dari studi tentang

database transaksi pelanggan untuk menentukan kebiasaan suatu produk dibeli bersama produk lainnya. Berdasarkan contoh, informasi dalam bentuk hubungan “jika-maka” atau “if-then” adalah informasi yang akan disampaikan oleh aturan asosiasi

$$\text{Confidence} = P(B|A) = \frac{\text{Jumlah T mengandung A dan B}}{\text{Jumlah Transaksi Mengandung A}}$$

Analisis asosiasi atau *Association Rule Mining* adalah teknik *Data Mining* untuk menemukan aturan asosiatif antara satu atau kombinasi item. Algoritma aturan asosiasi akan menggunakan data latihan, sesuai dengan pengertian *Data Mining*, untuk menghasilkan pengetahuan. Aturan asosiasi yang berbentuk “if...then...” atau “jika...maka...” merupakan pengetahuan yang dihasilkan dari fungsi aturan asosiasi (Fadlina, 2014). Metode *Association Rules* ini akan mendukung sistem rekomendasi penentuan kelulusan tepat waktu mahasiswa melalui data akademik mahasiswa.

2.4. Tahapan *Association Rules*

Metodologi *Association Rule* Terbagi menjadi dua tahap (Fadlina, 2014):
 1. Analisa pola frekuensi tinggi Tahap ini mencari kombinasi item yang memenuhi syarat minimum dari nilai support dalam database. Nilai support sebuah item diperoleh dengan rumus berikut:

$$\text{Support (A)} = \frac{\text{Jumlah Mengandung A}}{\text{Total Transaksi}}$$

Sedangkan nilai support dari 2 item diperoleh dari rumus berikut:

$$\text{Support (A,B)} = \frac{\text{Jumlah mengandung A dan B}}{\text{Total Transaksi}}$$

1. Pembentukan aturan Asosiatif

Setelah semua pola frekuensi tinggi ditemukan, barulah dicari aturan asosiatif yang memenuhi syarat minimum untuk *confidence* dengan menghitung *confidence* aturan asosiatif A_B Nilai *confidence* dari aturan A_B diperoleh dari rumus berikut:

2.5 Algoritma FP-Growth

Algoritma *Frequent Pattern-Growth* (*FP-Growth*) adalah salah satu cara alternatif untuk menemukan himpunan data yang paling sering muncul tanpa menggunakan generasi kandidat. *Frequent Pattern-Growth* (*FP-Growth*) membangun konstruksi data atau *Frequent Pattern Tree* (*FP-Tree*) yang sangat dikompresi, dan mengurangi data asli. Algoritma *Frequent Pattern Growth* (*FP-Growth*) melakukan scan database yang sama sebanyak dua kali. *Scanning database* yang pertama, kita dapat memperoleh *frequent 1-item-set*, dan *scanning database* yang kedua, kita dapat memfilter database *non-frequent item*, selebihnya, *Frequent Pattern Tree* (*FP-Tree*) dihasilkan secara bersamaan. Akhirnya, dapat diperoleh aturan asosiasi dengan menggunakan *Frequent Pattern Tree* (*FP-Tree*).

Sebuah *root* yang diberi label *null*, sekumpulan *sub-tree* yang beranggotakan *item-item* tertentu, dan sebuah tabel *frequent header*.

Setiap simpul dalam *Frequent Pattern Tree* (*FP-Tree*) mengandung tiga *field* yaitu (Muhammad Imam Ghazali, 2017):

- 1) *Item-name*: menginformasikan item yang dipresentasikan oleh simpul tersebut.
- 2) *Count*: mempresentasikan jumlah transaksi yang melewati simpul tersebut.
- 3) *Node-Link*: penghubung yang menghubungkan simpul-simpul dengan *item-name* yang sama, atau *null* jika kosong.

3. METODOLOGI PENELITIAN

Dalam melakukan penelitian ini ada beberapa Langkah-langkah Proses Perhitungan Asosiasi Proses Aturan Asosiasi terdiri dari beberapa tahap sebagai berikut (Tampubolo dan K. et.al., 2013).

- a. Sistem men-scan database untuk mendapatkandidat 1-itemset (himpunan *item* yang terdiri dari 1 *item*) dan menghitung nilai *support*nya. Kemudian nilai *Support*nya tersebut dibandingkan dengan *minimum support* yang telah ditentukan, jika nilainya lebih besar atau sama dengan *minimum support* maka *itemset* tersebut termasuk dalam *large itemset*.
- b. *Itemset* yang tidak termasuk dalam *large itemset* tidak diikuti dalam iterasi selanjutnya (di-prune).
- c. Pada iterasi kedua sistem akan menggunakan hasil *large itemset* pada iterasi pertama (L1) untuk membentuk kandidat *itemset* kedua (L2). Pada iterasi selanjutnya sistem akan menggunakan hasil *large itemset* pada iterasi selanjutnya akan menggunakan hasil *large itemset* pada iterasi sebelumnya (Lk-1) untuk membentuk kandidat *itemset* berikut (Lk). Sistem akan menggabungkan (*join*) Lk-1 dengan Lk-1 untuk mendapatkan Lk, seperti pada iterasi sebelumnya sistem akan menghapus (*prune*) kombinasi *itemset* yang tidak termasuk dalam *large itemset*.
- d. Setelah dilakukan operasi *join*, maka pasangan *itemset* baru hasil proses *join* tersebut dihitung *support*nya.
- e. Proses pembentuk kandidat yang terdiri dari proses *join* dan *prune* akan terus dilakukan hingga himpunan kandidat *itemset*nya *null*, atau sudah tidak ada lagi kandidat yang akan dibentuk.
- f. Setelah itu, dari hasil *frequent itemset* tersebut dibentuk *association rule* yang memenuhi nilai *support* dan *confidence* yang telah ditentukan.
- g. Pada pembentukan *association rule*, nilai yang sama dianggap sebagai satu nilai.
- h. *Assosiation rule* yang terbentuk harus memenuhi nilai minimum yang telah ditentukan.
- i. Untuk setiap *large itemset* L, kita cari himpunan bagian L yang tidak kosong. Untuk setiap himpunan bagian tersebut, dihasilkan *rule* dengan bentuk aB(L-a) jika *support* nya (L) dan *support* nya (a) lebih besar dari *minimum support*

4. HASIL DAN PEMBAHASAN

Data yang digunakan dalam proses *Data Mining* belum langsung bisa digunakan, dalam penelitian ini dilakukan transformasi data atau menggabungkan data agar proses *Mining* dapat digunakan langsung seperti tabel Tabel 4.1

Tabel 4.1: Data

Atribut	Keterangan
NIM	Nomor Induk Mahasiswa (NIM) adalah kode yang dimiliki mahasiswa sebagai nomor unik identitas diperguruan tinggi
IPK Semester 6	Indeks Prestasi Kumulatif (IPK) adalah Nilai mahasiswa dari semester 1-6
SKS Sampai semester 6	Merupakan jumlah SKS mahasiswa dari semester 1-6
Prediksi	Lulus Tepat Waktu/ Terlambat

Pada tabel 4.2 ada 1 dan 0 maksudnya adalah 1 adalah mahasiswa yang melakukan transaksi sedangkan 0 adalah mahasiswa yang tidak melakukan transaksi.

Tabel 4.2: Dataset

NO	NIM	B 2	A 2	A 3	A 1	B 1	B 3
1	13050038	1	0	1	0	0	0
2	13050044	1	0	1	0	0	0

3	1305004 5	1	1	0	0	0	0
4	1305011 4	0	0	1	0	0	1
5	1305005 1	1	0	1	0	0	0
6	1305002 1	1	1	0	0	0	0
7	1305000 1	1	1	0	0	0	0
8	1305003 7	0	0	1	0	0	1
9	1305012 2	1	0	1	0	0	0
10	1305005 8	0	0	1	0	0	1
11	1305010 1	0	0	1	0	0	1
12	1305008 2	1	0	1	0	0	0
13	1305010 5	1	1	0	0	0	0
14	1305010 7	1	0	0	1	0	0
15	1305000 2	1	0	1	0	0	0
16	1305013 0	1	0	1	0	0	0
17	1305006 8	1	1	0	0	0	0
18	1305005 6	1	1	0	0	0	0
19	1305012 3	1	1	0	0	0	0
20	1305002 3	1	1	0	0	0	0
21	1305002 2	0	0	0	1	1	0
22	1305003 4	1	1	0	0	0	0
23	1305007 0	1	1	0	0	0	0
24	1305007 8	1	1	0	0	0	0
25	1305005 2	1	1	0	0	0	0

26	1305008 9	1	1	0	0	0	0
27	1305006 0	1	1	0	0	0	0
28	1305009 6	0	1	0	0	1	0
29	1305012 1	1	1	0	0	0	0
30	1305002 4	1	0	0	1	0	0
32	1305009 2	1	0	0	1	0	0
33	1305007 6	0	0	0	1	1	0
34	1305004 2	0	0	0	1	1	0
35	1305008 1	0	0	0	1	1	0

Frekuensi Kemunculan setiap item dapat dilihat pada table 4.2 berikut :

Tabel 4.2 Frekuensi Kemunculan Tiap Item

ITEM	FREKUENS
	I
B2	25
A2	16
A3	11
A1	8
B1	6
B3	4

2.Menentukan Minumun Support
 Berikut ini Merupakan Hasil Pemindaian yang memiliki frekuensi di atas *support count* $\xi = 5$ adalah B2,A2,A3,A1 dan B1.

Tabel 4.3 Frequent list

ITEM	FREKUENSI
B2	25
A2	16
A3	11
A1	8
B1	6

Pada Tabel 4.3 Sesudah pemindaian

dilakukan maka didapatkan *item* yang memiliki frekuensi di atas *support count* $\xi = 5$. IPK dengan rentang 2,76-3,50 (B2), Total SKS 129-136 SKS (A2), Total SKS 137-142 SKS (A3), Total SKS 116-126 (A1), IPK dengan rentang 2,00-2,75 (B1). Dari 5 *item* inilah yang nantinya akan berpengaruh dan juga akan dimasukkan ke dalam *Fp-tree*,

3..Menentukan *Header Frequent Itemset*

Tabel 4.4 Data transaksi

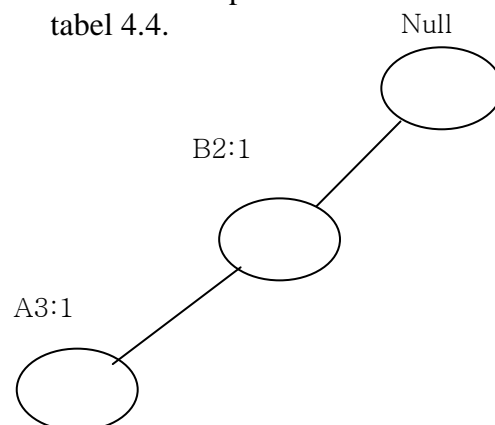
TID	ITEMSET
13050038	{B2, A3}
13050044	{B2, A3}
13050045	{B2, A2}
13050114	{A3}
13050051	{B2, A3}
13050021	{B2, A2}
13050001	{B2, A2}
13050037	{A3}
13050122	{B2, A3}
13050058	{A3}
13050101	{A3}
13050082	{B2, A3}
13050105	{B2, A2}
13050107	{B2, A1}
13050002	{B2, A3}
13050130	{B2, A3}
13050068	{B2, A2}
13050056	{B2, A2}
13050123	{B2, A2}
13050023	{B2, A2}
13050022	{B1, A1}
13050034	{B2, A2}
13050070	{B2, A2}
13050078	{B2, A2}
13050052	{B2, A2}
13050089	{B2, A2}
13050060	{B2, A2}
13050096	{B1, A2}
13050121	{B2, A2}

13050024	{B2, A1}
13050057	{B1, A1}
13050092	{B2, A1}
13050076	{B1, A1}
13050042	{B1, A1}
13050081	{B1, A1}

4.Menentukan *Fp-Tree*

Langkah berikut ini merupakan suatu pembentukan *Fp-Tree*.

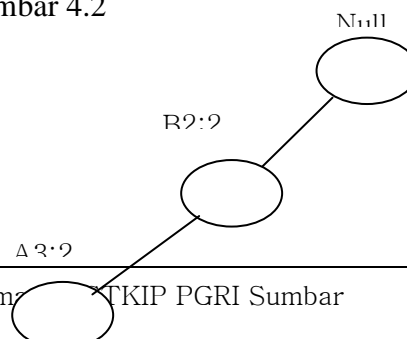
- a. Berikut gambar yang memberikan ilustrasi mengenai pembentukan *Fp-Tree* setelah pembacaan TID 1 pada tabel 4.4.



Gambar 4.1 Hasil Pembentukan TID 1

Gambar 4.1 adalah penjelasan tentang pembentukan *Fp-Tree* setelah pembacaan didapat setelah melakukan TID 1, yaitu berisi Null-B2=1 (IPK dengan rentang 2,76-3.50) dan A3=1 (SKS 137-142).

- b. Pembentukan *Fp-tree* ke dua pada gambar 4.2

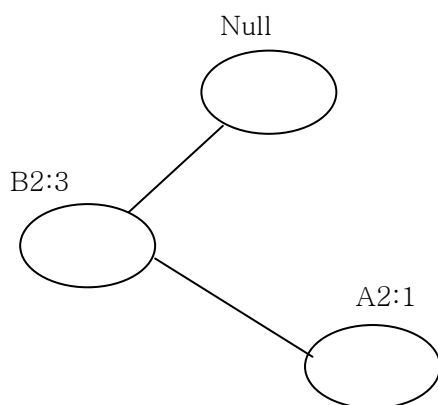


Gambar 4.4 Hasil Pembentukan TID 4

Gambar 4.2 Hasil Pembentukan TID 2

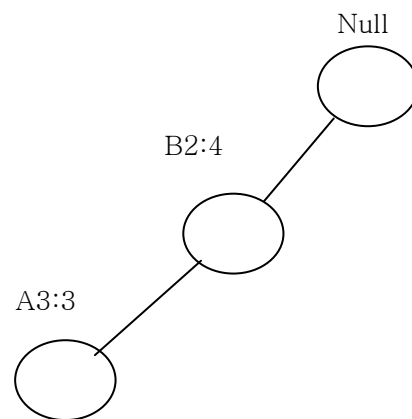
Gambar 4.2 adalah penjelasan tentang pembentukan *Fp-Tree* setelah pembacaan didapat setelah melakukan TID 2, yaitu berisi Null-B2=2 (IPK dengan rentang 2,76-3.50) dan A3=2 (SKS137-142).

- c. Pembentukan *Fp-tree* dari tindakan ke tiga pada gambar 4.3:



Gambar 4.4 adalah penjelasan tentang pembentukan *Fp-Tree* setelah pembacaan didapat setelah melakukan TID 4, yaitu berisi Null-A3=1 (Total SKS 137-142 SKS)

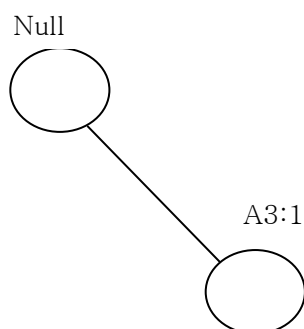
- e. Pembentukan *Fp-tree* dari tindakan ke lima pada gambar 4.5:



Gambar 4.3 Hasil Pembentukan TID 3

Gambar 4.3 adalah penjelasan tentang pembentukan *Fp-Tree* setelah pembacaan didapat setelah melakukan TID 3, yaitu berisi Null-B2=3 (IPK dengan rentang 2,75-3,50) dan A2=1 (Total SKS127-136 SKS).

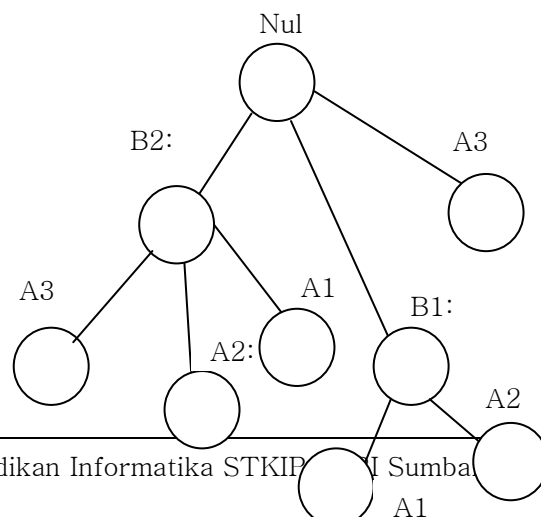
- d. Pembentukan *Fp-tree* dari tindakan ke empat pada gambar 4.4:



Gambar 4.5 Hasil Pembentukan TID 5

Gambar 4.5 adalah penjelasan tentang pembentukan *Fp-Tree* setelah pembacaan didapat setelah melakukan TID 5, yaitu berisi Null-B2=4 (IPK dengan rentang 2,75-3,50) dan A3=3 (Total SKS 137-142 SKS).

- f. Pembentukan *Fp-tree* dari tindakan ke 35 pada gambar 4.6:

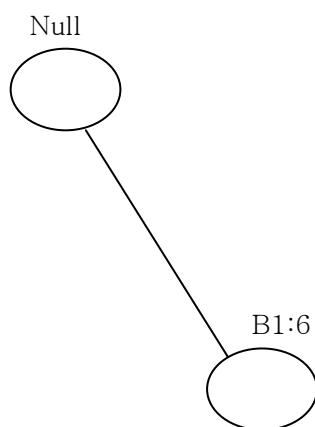


Gambar 4.6 Hasil Pembentukan TID 35

Gambar 4.6 adalah penjelasan tentang pembentukan *Fp-Tree* setelah pembacaan didapat setelah melakukan TID 35 yaitu berisi Null- B2=25 (IPK dengan rentang 2,76-3.50), A2=16 (SKS 127-136 SKS), A3=11 (SKS dengan rentang 137-142 sks), A1=8 (SKS 116-126), B1=6 (IPK dengan rentang 2,00-2,75).

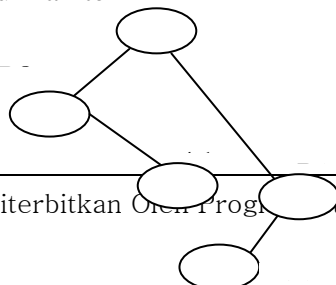
5. Menentukan *frequent Itemset*

Untuk menemukan *Frequent item set* dari tabel 4.2, maka perlu ditentukan terlebih dahulu lintasan yang berakhir dengan *support count* terkecil, yaitu B1 yang diikuti dengan A1, A3, A2 dan diakhiri B2. Proses pembentukan masing-masing node dapat dilihat pada gambar 4.7 Berikut.



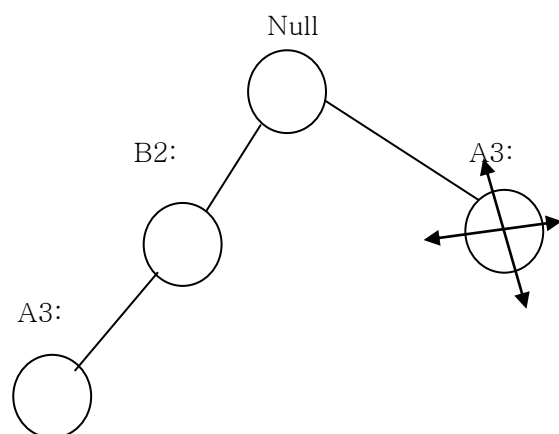
Gambar 4.7 Lintasan Yang Mengandung B1

Gambar 4.7 adalah penjelasan tentang lintasan yang berakhir B1 yang memiliki jumlah ter



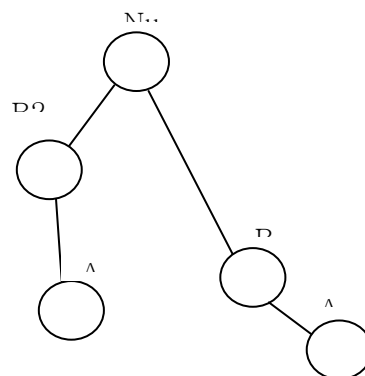
Gambar 4.8 Lintasan Yang Mengandung A1

Gambar 4.8 adalah penjelasan tentang lintasan yang berakhir A1 yang memiliki frekuensi kemunculan sebanyak 8 kali



Gambar 4.9 Lintasan Yang Mengandung A3

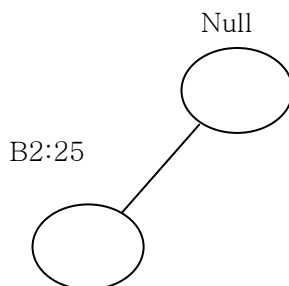
Gambar 4.9 adalah penjelasan tentang lintasan yang berakhir A3 yang memiliki frekuensi kemunculan sebanyak 11 kali tahap. Gambar 4.9 mempertlihatkan nilai item pada *tree* yang kurang dari batas nilai ambang atau nilai *threshold* akan dihapus.



Gambar 4.10 Lintasan Yang Mengandung

A2

Gambar 4.10 adalah penjelasan tentang lintasan yang berakhiran A2 yang memiliki frekuensi kemunculan sebanyak 16 kali.



Gambar 4.11 Lintasan Yang Mengandung B2

Gambar 4.11 adalah penjelasan tentang lintasan yang berakhiran B2 yang memiliki frekuensi kemunculan sebanyak 25 kali.

6. Membuat Conditional Pattern

Setelah mencari *frequent itemset* untuk beberapa akhiran *suffix* maka didapat hasil yang dirangkum dalam Tabel 4.6:

Tabel 4.5

Suffix	Frekuensi Itemset
B1	{B1}
A1	{B1,A1}
A3	{B2,A3}
A2	{B2,A2}
B2	{B2}

7. Menghitung Support dan Confidence

Selanjutnya menghitung *support* dan *confidence* dari hasil pembentukan *fp-tree* yang sudah dilakukan dengan menggunakan rumus.

Dengan rumus (1)

$$Support(A) = \frac{5}{35} \cdot 100\% = 14,2\%$$

Untuk Σ Item pada kasus { *If* IPK dengan rentang 2,76-3.50 (B2) then SKS 127-136

SKS (A2), ada 5 dari 35 kasus sehingga untuk *Support* nya adalah 14,2 %.

Dengan rumus (2)

$$Confidence = P(B|A) = \frac{5}{15} \cdot 100\% = 33\%$$

Untuk Σ Item pada kasus { *If* IPK dengan rentang 2,76-3.50 (B2) then SKS 127-136 SKS (A2) ada 5 dari 15 orang sehingga untuk *Confidence* nya adalah 33%.

Tabel 4.6 Tabel *Support* dan *Confidence* 2 Item

<i>If Antecedent then Consequent</i>	<i>Support</i>	<i>Confidence</i>
<i>If</i> IPK dengan rentang 2,76-3.50 (B2) <i>Then</i> SKS 127-136 SKS (A2)	15/35=42,8%	15/25=60%
<i>If</i> IPK dengan rentang 2,76-3.50 (B2) <i>Then</i> SKS 137-142 SKS (A3)	7/35=20%	7/25=28%
<i>If</i> IPK dengan rentang 2,00-2,75 (B1) <i>Then</i> SKS 116 SKS - 126 SKS (A1)	5/35=14%	5/6=83%
<i>If</i> SKS 137-142 SKS (A3) <i>Then</i> IPK dengan rentang 2,76-3.50 (B2)	7/35=20%	7/11=63,6%
<i>If</i> SKS 127-136 SKS (A2) <i>Then</i> IPK dengan rentang 2,76-3.50 (B2)	15/35=42,8%	15/16=93,7%
<i>If</i> SKS 116 SKS - 126 SKS (A1) <i>Then</i> IPK dengan rentang 2,00-2,75 (B1)	5/35=14%	5/8=62,5%

Dari *rule* yang didapat, berikut ini merupakan *rule* yang memiliki *Support* $\geq 14,2\%$ dan memiliki *Confidence* $\geq 33\%$ seperti pada tabel 4.6

Padabagianberisipenjelasanilmiahdarihasilpen erapanmetodepenelitian yang telahditetapkanpada sub bab 3.

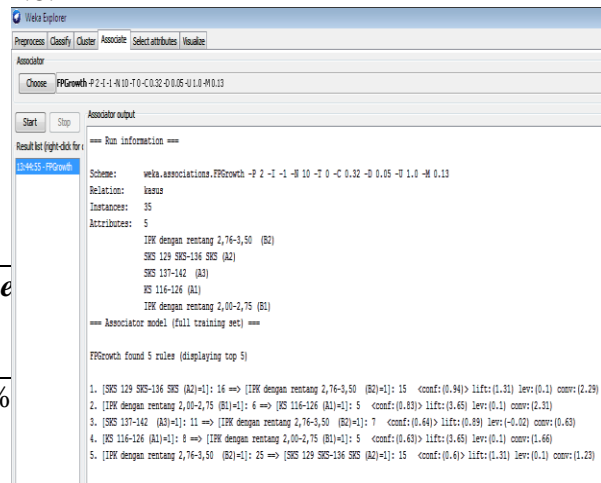
6. Hasil Rule

Tabel 4.7: Hasil Rule

If Antecedent then Consequent	Support	Confidence
If IPK dengan rentang 2,76-3.50 (B2) Then SKS 127-136 SKS (A2)	15/35=42,8%	15/25=60%
If IPK dengan rentang 2,00-2,75 (B1) Then SKS 116 SKS - 126 SKS (A1)	5/35 =14%	5/6= 83%
If SKS 137-142 SKS (A3) Then IPK dengan rentang 2,76-3.50 (B2)	7/35=20%	7/11=63,6%
If SKS 127-136 SKS (A2) Then IPK dengan rentang 2,76-3.50 (B2)	15/35=42,8%	15/16=93,7%
If SKS 116 SKS -126 SKS (A1) Then IPK dengan rentang 2,00-2,75 (B1)	5/35 =14%	5/8=62,5%

Pada tabel 4.7 didapatkan *rule* yang mempunyai *Confidence* $\geq 100\%$ kemudian *rule* yang memenuhi kebutuhan, maka dapat disimpulkan *rule* yang terbaik adalah **If SKS 127-136 SKS (A2) Then IPK dengan rentang 2,76-3.50 (B2)**. Dari hasil *rule* terbaik tersebut menjadi patokan untuk

mengetahui kelulusan tepat waktu mahasiswa. Untuk lebih jelasnya bias dilihat pada table 4.8.



KESIMPULAN

Bagian terdiri atas simpulan dan saran atas penelitian hasil penelitian.

5.1 Simpulan

Setelah dilakukan pengujian pada program atau menggunakan aplikasi WEKA maka dapat diambil kesimpulan bahwa:

1. Sistem yang baru dapat diterapkan sehingga Membantu Program Studi Pendidikan Matematika STKIP PGRI Sumatera Barat mengetahui tingkat kelulusan tepat waktu mahasiswa menggunakan metode *Association Rule* dengan menggunakan algoritma *FP-Growth* berdasarkan IPK Semester 6 dan Total SKS di Semester 6.
2. Analisa dan pertimbangan terhadap hasil asosiasi mahasiswa lulus tepat waktu menggunakan algoritma *FP-Growth* dapat diketahui dari asosiasi yang memiliki persentase tertinggi dari IPK Semester 6 dan Total SKS di Semester 6.
3. Aplikasi dapat Membantu Program Studi Pendidikan Matematika STKIP PGRI Sumatera Barat mengetahui tingkat kelulusan tepat waktu

mahasiswa menggunakan Aplikasi Weka.

5.2 Saran

Setelah melakukan penelitian ini, penulis ingin menyampaikan beberapa saran yang bermanfaat untuk kesempurnaan penelitian ini :

1. Penulis menyarankan adanya pengembangan *Data Mining Association Rule* dengan Algoritma *FP-Growth* maka pertimbangan dan analisa tidak perlu dilakukan secara manual, tetapi dilakukan secara otomatis
2. Penulis menyarankan adanya pengembangan lebih lanjut dengan menggunakan teknik *asosiasi* menggunakan algoritma selain *FP-Growth* serta dapat menggunakan teknik *Data Mining* yang lain. Sehingga menambah ilmu pengetahuan.

DaftarRujukan

- Tri Ferga Prasetyo, Dony Susandi, Ida Siti Widianingrum**, 2016. Prediksi Kelulusan Mahasiswa Pada Perguruan Tinggi Kabupaten Majalengka Berbasis *Knowledge Based System*. Bandung : **Teknik Informatika**.
- Muhammad Imam Ghozali**, 2017. Analisa Pola Belanja Menggunakan Algoritma *FP-Growth*, *Self Organizing Map* (SOM) dan *K-Medoids*. **Jurnal Simetris**.
- Haroe Snatoso, I Putu Hariyadi, Prayitno** , 2016. *Data Mining Analisa Pola Pembelian Produk dengan Menggunakan Metode Algoritma Apriori*. Yogyakarta: **Teknik Informatik**

a

- K. Sumathi, S. Kannan, K. Nagarajan**, 2016. *Data Mining: Analysis Of Studend Database Using Classification Techniques. International Journal of Computer Application, Volume 4*
- Kennedi Tampubolon, Hoga Saragih, Bobby Reza**, 2013. Implementasi *Data Mining* Algoritma Apriori Pada Sistem Persediaan Alat-alat Kesehatan. **Informasi dan Teknologi IImiah. Volume 1**
- M. Durairaj, V. Ranjani** , 2013. *Data Mining Applications In Healthcare Sector : A Study. International Journal Of Scientific & Technology Researctch Volume 2*.
- David Hartanto Kamagi, Seng Hansun** , 2014. Implementasi *Data Mining* dengan Algoritma C4.5 untuk Memprediksi Tingkat Kelulusan Tepat Waktu Mahasiswa. **ULTIMATIC, Vol. VI, No.1**
- Hermansyah Nur Ahmad, Vincent Suhartono, Ika Novita Dewi** ,2017. Penentuan Tingkat Kelulusan Tepat Waktu Mahasiswa STMIK Subang Menggunakan Algoritma C4.5. **Jurnal Teknologi Informasi, Volume 13 Nomor 1**.
- Priyadharsini C & Thanamani Antony Selvadoss Thanamani**, 2014. An Overview of Knowledge Discovery Database and Data Mining Techniques. **Departement of CSE., JayShriram Group of Institutions, Tirupur, Tamilnadu, India on 6th & 7th March 2014**
- Fadlina**, 2014. *Data Mining untuk Analisa Tingkat Kejahatan Jalanan dengan Algoritma Association Rule Metode Apriori*. **UPI YPTK Padang**.